REMARKS

Claims 7-11 were pending in the instant application. The specification has been amended to include a table number. Claims 8-10 have been cancelled without prejudice. Claims 7 and 11 have been amended in order to claim more fully and distinctly the invention. Accordingly, claims 7 and 11 are currently pending in the present application. Support for the amended table number and claims can be found throughout the specification and claims as originally filed. Specifically, support for the amended table number may be found at least at page 48, line 7 and page 55, line 4. Support for the amended claim 7 may be found at least, for example, at page 7, line 30 through page 8, line 3. Support for the amended claim 11 may be found at least, for example, at page 5, lines 12-23. No new matter has been added.

Attached hereto is Appendix A, captioned "VERSION WITH MARKINGS TO SHOW CHANGES MADE". The attached Appendix includes a marked-up version of the changes made to the specification by the current amendment. Also attached hereto is APPENDIX B, including the full set of claims that are currently pending. Also attached hereto is APPENDIX C, including the SEQ ID NO:2 from U.S. Serial No. 60/044,185, filed March 24, 1997. Also attached hereto is APPENDIX D, including the SEQ ID NO:15 from U.S. Serial No. 09/062,142, filed March 17, 1998.

Amendment of the claims is not to be construed as an acquiescence to any of the rejections set forth in the instant Office Action, and was done solely to expedite prosecution of the instant application. Applicants reserve the right to pursue the claims as originally filed, or similar claims, in this or one or more patent applications.

Claim Rejections Under 35 U.S.C. §101

Claims 7-11 stand rejected under 35 U.S.C. §101 on the ground that the claimed invention is not supported by a specific asserted utility or a substantial utility. Applicants respectfully disagree and traverse the foregoing rejection for the following reasons.

Claims 8-10 have been cancelled, rendering the rejection moot with respect to these claims. Moreover, with respect to the presently pending claims 7 and 11, it is Applicants' position that a specific and substantial utility for the claimed invention is clearly set forth in the instant specification and the knowledge in the art at the time of Applicants' invention.

Applicants disclose in the instant specification a full-length cDNA which contains an open reading frame encoding the polypeptide of SEQ ID NO:3. Applicants assert in the instant specification that the protein has a significant degree of homology to trypsin-type serine proteases (see, *e.g.* page 50, lines 1-3). Moreover, functional analysis of this protein demonstrated urokinase activity upon transfection into COS7 cells (see, *e.g.* page 49, line 23-25), indicating that this protein shares the activities of urokinase-like proteins.

Urokinase-like proteins are a subgroup of the serine protease family of enzymes and include, but are not limited to, urokinase, streptokinase, vascular plasminogen activator and tissue plasminogen activator. These enzymes are well-known in the art by their mechanism of action, which is based on the formation of an acyl enzyme intermediate on a specific active serine residue. Specifically, urokinase-like molecules act as plasminogen activators, acting on plasminogen to generate plasmin. This activity is important for a number of biological functions, including wound healing, hemostasis and thrombolysis.

Applicants assert that novel molecules of the present invention can be used, for example, as modulators of hemostatic and thrombolytic activity, i.e. for dissolving or inhibiting formation of thromboses (see, *e.g.* page 33, lines 19-23) or modulating coagulation (see, *e.g.*, page 33, lines 13-19), or as modulators of tissue growth activity, i.e. for use in wound healing (see, *e.g.*, page 27, lines 6-8). Accordingly, the polypeptide of the present invention can be used for diagnostic and therapeutic purposes for disorders which involve any of these biological activities (see, *e.g.*, the specification, at least, for example, at page 29, lines 22-26 and page 33, lines 13-23).

The specificity of the asserted utilities is based on the fact that the polypeptide of the present invention belongs to the urokinase-like protein subfamily of serine proteases, a family sharing structural and functional characteristics which are not shared by other non-urokinase proteins. In particular, urokinases are known to act as plasminogen activators, cleaving plasminogen to generate plasmin. This activity is important for a number of biological functions, including wound healing, hemostasis and thrombolysis. Applicants respectively assert that these activities are specific to the urokinase subfamily of serine proteases and are not shared by all other protein-encoded nucleic acid molecules.

Moreover, no evidence has been made of record that Applicants' assertions regarding the activity and/or utility of SEQ ID NO:3 polypeptides as modulators of tissue growth activity, i.e.

wound healing, or hemostatic and/or thrombolytic activity would not be considered credible to one of skill in the art. As the Examiner is aware, an applicant must provide only one credible assertion of utility for any claimed invention to satisfy the utility requirement. The instant application teaches a specific and substantial biological function for the SEQ ID NO:3 polypeptides of the invention. Accordingly, Applicants respectfully request reconsideration and withdrawal of the rejection of the claims under 35 U.S.C. §101.

Claim Rejections Under 35 U.S.C. §112, first paragraph

Claims were rejected 7-11 under 35 U.S.C. §112, first paragraph. Specifically, the Office Action states that "[s]ince the claimed invention is not supported by either a specific and substantial utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention." Applicants respectfully traverse.

Without acquiescing to the alleged lack of enablement in the specification as originally filed, claims 8-10 have been cancelled, thus rendering the rejection moot. Applicants submit that the rejection with regard to these claims is therefore obviated.

With respect to newly amended claims 7 and 11, Applicants would like to make the following remarks of record. As argued above, the present invention is supported by a substantial utility and a well-established utility. Specifically, the asserted utilities are based on the fact that SEQ ID NO:3 polypeptides of the present invention are urokinase-like proteins, a subfamily of the serine protease enzyme family. Further, the specification is replete with teachings of how to make and/or use the present invention. For example, the specification teaches that novel molecules of the present invention can be used, for example, as modulators of hemostatic and thrombolytic activity, i.e. for dissolving or inhibiting formation of thromboses (see, *e.g.* page 33, lines 19-23) or modulating coagulation (see, *e.g.*, page 33, lines 13-19), or as modulators of tissue growth activity, i.e. for use in wound healing (see, *e.g.*, page 27, lines 6-8). Accordingly, the polypeptide of the present invention can be used for diagnostic and therapeutic purposes for disorders which involve any of these biological activities (see, *e.g.*, the specification, at least, for example, at page 29, lines 22-26 and page 33, lines 13-23). Applicants respectfully submit that any experimentation that may be required to make and/or use the

claimed polypeptide molecules constitutes routine, not undue, experimentation and therefore the specification clearly enables the pending claims.

Claims 8-11 were further rejected under 35 U.S.C. §112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention.

Without acquiescing to the alleged lack of written description in the specification as originally filed, claims 8-10 have been cancelled, thus rendering the rejection moot as it applies to these claims. Claim 11 has been amended to depend from claim 7, thus rendering the rejection moot as it applies to this claim. Applicants submit that the rejection with regard to these claims is therefore obviated.

Claims 8-10 were also rejected under 35 U.S.C. §112, first paragraph, for lack of enablement. Specifically, the Office Action states that while the specification is enabling for a protein encoded by SEQ ID NO:3, it does not enable a "polypeptide consisting of a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:3" or "a polypeptide consisting of a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:3" or "a polypeptide which is at least 60% homologous to a polypeptide encoded by the amino acid sequence of SEQ ID NO:3."

Without acquiescing to the alleged lack of enablement in the specification as originally filed, claims 8-10 have been cancelled, thus rendering the rejection moot. Applicants submit that the rejection with regard to these claims is therefore obviated.

Claim Rejections Under 35 U.S.C. §112, second paragraph

Claims 8, 9 and 11 were rejected under 35 U.S.C. 35 §112, second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which Applicant regards as the invention. Specifically, the Office Action states that claim 8 is rejected as vague and indefinite in the recitation of the phrase "fragment," claim 11 is rejected insofar as it depends on claim 8, and claim 9 is rejected as vague and indefinite for reciting the phrases "allelic variant" and "stringent conditions."

Without acquiescing to the alleged lack of definiteness in the specification as originally filed, claims 8 and 9 have been cancelled, thus rendering the rejection moot as it applies to these

claims. Claim 11 has been amended to depend from claim 7, thus rendering the rejection moot as it applies to this claim. Applicants submit that the rejection with regard to these claims is therefore obviated.

Claim Rejections Under 35 U.S.C. 102(e)

Claims 7-11 were rejected under 35 U.S.C. §102(e) as being anticipated by Sheppard (U.S. Patent No. 6,153,420). Specifically, the Office Action states that "Sheppard teaches a serine protease polypeptides [sic]. The polypeptide sequence (SEQ ID NO:18) described by Sheppard in U.S. Patent No. 6,153,420 (columns 41-44) has over 100% identity over its entire length to SEQ ID NO:3 of the instant invention. Therefore, the disclosure of Sheppard anticipates claims 7-11." Applicants respectfully traverse.

Applicants respectfully submit that the cancellation of claims 8-10, without prejudice, renders the aforementioned rejection moot and request that the Examiner withdraw this §102(e) rejection as it pertains to these claims.

With respect to newly amended claims 7 and 11, Applicants submit the following remarks for the record. The polypeptide sequence (SEQ ID NO:18) was described by Sheppard in U.S. Patent No. 6,153,420 ('420), filed May 4, 1998, which claims priority to U.S. Serial No. 09/062,142, filed March 17, 1998 and U.S. Serial No. 60/044,185, filed April 24, 1997. However, Applicants submit that SEQ ID NO:18 as described in the '420 patent is not entitled to these earlier priority dates.

First, U.S. Serial No. 60/044,185 discloses SEQ ID NO:2, which differs from the disclosed SEQ ID NO:18 of the '420 patent and SEQ ID NO:3 of the present invention in that it contains 2 different amino acids at positions 60 and 299, 4 unidentified amino acids at positions 80, 95, 96 and 149 as well as 9 additional amino acids at the C-terminal end (amino acids 365-373) (See attached Appendix C). Second, U.S. Serial No. 09/062,142 discloses SEQ ID NO:15, which differs from the disclosed SEQ ID NO:18 of the '420 patent and SEQ ID NO:3 of the present invention in that it also contains 9 additional amino acids at the C-terminal end of the polypeptide (amino acids 365-373) (See attached Appendix D). Therefore, because SEQ ID NO:18 is different from those sequences found in the applications from which priority is

claimed, Applicants respectfully submit that SEQ ID NO:18 is entitled to only the May 4, 1998 filing date of the '420 patent.

Applicants first described and disclosed the polypeptide SEQ ID NO:3 in the Japanese Patent Application JP 9/323129, filed November 25, 1997, which is ***before*** the May 4, 1998 priority date of the '420 patent by Sheppard. Therefore, Sheppard is unavailable as prior art against the instant application as it was filed ***after*** the priority date if the instant application. In view of the above, Applicants respectfully request that the Examiner withdraw the rejection of claims 7-11 under 35 U.S.C. §102(e).

CONCLUSION

If a telephone conversation with Applicants' attorney would help expedite the prosecution of the above-identified application, the Examiner is urged to call Applicants' attorney at (617) 227-7400.

Respectfully submitted,
LAHIVE & COCKFIELD, LLP

Cynthia L. Kanik, Ph.D.
Reg. No. 37,320
Attorney for Applicants

28 State Street
Boston, MA 02109
Tel: (617) 227-7400
Fax: (617) 742-4214

Date: **February 8, 2002**

APPENDIX A

Version With Markings To Show Changes Made

In the specification:

The sentence on page 55, line 4 has been replaced with the following rewritten sentence:

-- Table 4--

In the claims:

Claims 7 and 11 have been amended as follows:

7. (Amended) An isolated polypeptide ~~comprising~~ consisting of the amino acid sequence of SEQ ID NO:3.

11. (Amended) The isolated polypeptide of claim ~~8~~ 7, further comprising heterologous amino acid sequences.

APPENDIX B

Pending Claims

7. (Amended) An isolated polypeptide consisting of the amino acid sequence of SEQ ID NO:3.

11. (Amended) The isolated polypeptide of claim 7, further consisting of heterologous amino acid sequences.

APPENDIX C

FILE WRAPPER FOR PROVISIONAL U.S. APPLICATION

NO: **60/044,185**

INVENTOR: PAUL O. SHEPPARD
LAURA JELINEK
DONALD C. FOSTER

FILING DATE: MARCH 24, 1997

TITLE: SERINE PROTEASE POLYPEPTIDES AND MATERIALS AND METHODS FOR MAKING THEM

---

*RELATED U.S. APPLICATION DATA:

USSN 09/072,384 FILED MAY 4, 1998
US PATENT 6,153,420

US PROVISIONAL APPLICATION NO. 60/044,185
FILED APRIL 24, 1997 **[Captioned file]**

USSN 09/062,142 FILED APRIL 17, 1998
ABANDONED

*The related U.S. application data is drawn from the USPTO's public website and is not to be construed as a complete family of applications. Complete family information is available from the USPTO under 37 CFR §1.14.

AGAGATATG 1634

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 392 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(ix) FEATURE:

(A) NAME/KEY: Signal Sequence
(B) LOCATION: 1...19
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Gly Ile Pro Gly Leu Leu Phe Leu Leu Phe Phe Leu Leu Cys
                -15                 -10                 -5
Ala Val Gly Gln Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr Trp
            1               5                  10
Pro Ala Tyr Arg Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu
    15                  20                  25
Ala Lys Pro Asp Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser
30                  35                  40                  45
Cys Gly Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Lys Glu
                50                  55                  60
Ala Lys Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg
            65                  70                  75
Thr Glu Xaa Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly
        80                  85                  90
Ala Xaa Xaa Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg
    95                  100                 105
Gln Ile Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe
110                 115                 120                 125
Leu Leu Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys
                130                 135                 140
Thr Gly Thr Leu Val Ala Glu Xaa His Val Leu Thr Ala Ala His Cys
            145                 150                 155
Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val
        160                 165                 170
```

```
Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp
    175                 180                 185
Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val
190                 195                 200                 205
Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp
                210                 215                 220
Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His
            225                 230                 235
Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu
        240                 245                 250
Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly
    255                 260                 265
Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu
270                 275                 280                 285
Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Tyr Gly Val
                290                 295                 300
Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile
            305                 310                 315
Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro
        320                 325                 330
Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln
    335                 340                 345
Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp
350                 355                 360                 365
Thr Val Phe Leu Pro Gly Ser Asn
                370
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGYACNGGNW SNHTNRT 17

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid

APPENDIX D

FILE WRAPPER FOR ABANDONED U.S. APPLICATION

SERIAL NO: **09/062,142**

INVENTORS: PAUL O. SHEPPARD

FILING DATE: MARCH 17, 1998

TITLE: SERINE PROTEASE POLYPEPTIDES AND MATERIALS AND METHODS FOR MAKING THEM

---

*RELATED U.S. APPLICATION DATA:

USSN 09/072,384 FILED MAY 4, 1998
US PATENT 6,153,420

US PROVISIONAL APPLICATION NO. 60/044,185
FILED APRIL 24, 1997

USSN 09/062,142 FILED APRIL 17, 1998
ABANDONED **[Captioned file]**

*The related U.S. application data is drawn from the USPTO's public website and is not to be construed as a complete family of applications. Complete family information is available from the USPTO under 37 CFR §1.14.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Ala Gly Ile Pro Gly Leu Leu Phe Leu Leu Phe Phe Leu Leu Cys
                -15                 -10                 -5
Ala Val Gly Gln Val Ser Pro Tyr Ser Ala Pro Trp Lys Pro Thr Trp
            1               5                  10
Pro Ala Tyr Arg Leu Pro Val Val Leu Pro Gln Ser Thr Leu Asn Leu
    15                  20                  25
Ala Lys Pro Asp Phe Gly Ala Glu Ala Lys Leu Glu Val Ser Ser Ser
30                  35                  40                  45
Cys Gly Pro Gln Cys His Lys Gly Thr Pro Leu Pro Thr Tyr Glu Glu
                50                  55                  60
Ala Lys Gln Tyr Leu Ser Tyr Glu Thr Leu Tyr Ala Asn Gly Ser Arg
            65                  70                  75
Thr Glu Thr Gln Val Gly Ile Tyr Ile Leu Ser Ser Ser Gly Asp Gly
        80                  85                  90
Ala Gln His Arg Asp Ser Gly Ser Ser Gly Lys Ser Arg Arg Lys Arg
    95                  100                 105
Gln Ile Tyr Gly Tyr Asp Ser Arg Phe Ser Ile Phe Gly Lys Asp Phe
110                 115                 120                 125
Leu Leu Asn Tyr Pro Phe Ser Thr Ser Val Lys Leu Ser Thr Gly Cys
                130                 135                 140
Thr Gly Thr Leu Val Ala Glu Lys His Val Leu Thr Ala Ala His Cys
            145                 150                 155
Ile His Asp Gly Lys Thr Tyr Val Lys Gly Thr Gln Lys Leu Arg Val
        160                 165                 170
Gly Phe Leu Lys Pro Lys Phe Lys Asp Gly Gly Arg Gly Ala Asn Asp
    175                 180                 185
Ser Thr Ser Ala Met Pro Glu Gln Met Lys Phe Gln Trp Ile Arg Val
190                 195                 200                 205
Lys Arg Thr His Val Pro Lys Gly Trp Ile Lys Gly Asn Ala Asn Asp
                210                 215                 220
Ile Gly Met Asp Tyr Asp Tyr Ala Leu Leu Glu Leu Lys Lys Pro His
            225                 230                 235
Lys Arg Lys Phe Met Lys Ile Gly Val Ser Pro Pro Ala Lys Gln Leu
        240                 245                 250
Pro Gly Gly Arg Ile His Phe Ser Gly Tyr Asp Asn Asp Arg Pro Gly
    255                 260                 265
Asn Leu Val Tyr Arg Phe Cys Asp Val Lys Asp Glu Thr Tyr Asp Leu
270                 275                 280                 285
Leu Tyr Gln Gln Cys Asp Ala Gln Pro Gly Ala Ser Gly Ser Gly Val
                290                 295                 300
Tyr Val Arg Met Trp Lys Arg Gln Gln Gln Lys Trp Glu Arg Lys Ile
            305                 310                 315
```

```
Ile Gly Ile Phe Ser Gly His Gln Trp Val Asp Met Asn Gly Ser Pro
        320                 325                 330
Gln Asp Phe Asn Val Ala Val Arg Ile Thr Pro Leu Lys Tyr Ala Gln
    335                 340                 345
Ile Cys Tyr Trp Ile Lys Gly Asn Tyr Leu Asp Cys Arg Glu Gly Asp
350                 355                 360                 365
Thr Val Phe Pro Pro Gly Ser Asn
                370
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1176 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
ATGGCNGGNA THCCNGGNYT NYTNTTYYTN YTNTTYTTYY TNYTNTGYGC NGTNGGNCAR     60
GTNWSNCCNT AYWSNGCNCC NTGGAARCCN ACNTGGCCNG CNTAYMGNYT NCCNGTNGTN    120
YTNCCNCARW SNACNYTNAA YYTNGCNAAR CCNGAYTTYG GNGCNGARGC NAARYTNGAR    180
GTNWSNWSNW SNTGYGGNCC NCARTGYCAY AARGGNACNC CNYTNCCNAC NTAYGARGAR    240
GCNAARCART AYYTNWSNTA YGARACNYTN TAYGCNAAYG GNWSNMGNAC NGARACNCAR    300
GTNGGNATHT AYATHYTNWS NWSNWSNGGN GAYGGNGCNC ARCAYMGNGA YWSNGGNWSN    360
WSNGGNAARW SNMGNMGNAA RMGNCARATH TAYGGNTAYG AYWSNMGNTT YWSNATHTTY    420
GGNAARGAYT TYYTNYTNAA YTAYCCNTTY WSNACNWSNG TNAARYTNWS NACNGGNTGY    480
ACNGGNACNY TNGTNGCNGA RAARCAYGTN YTNACNGCNG CNCAYTGYAT HCAYGAYGGN    540
AARACNTAYG TNAARGGNAC NCARAARYTN MGNGTNGGNT TYYTNAARCC NAARTTYAAR    600
GAYGGNGGNM GNGGNGCNAA YGAYWSNACN WSNGCNATGC CNGARCARAT GAARTTYCAR    660
TGGATHMGNG TNAARMGNAC NCAYGTNCCN AARGGNTGGA THAARGGNAA YGCNAAYGAY    720
ATHGGNATGG AYTAYGAYTA YGCNYTNYTN GARYTNAARA ARCCNCAYAA RMGNAARTTY    780
ATGAARATHG GNGTNWSNCC NCCNGCNAAR CARYTNCCNG GNGGNMGNAT HCAYTTYWSN    840
GGNTAYGAYA AYGAYMGNCC NGGNAAYYTN GTNTAYMGNT TYTGYGAYGT NAARGAYGAR    900
ACNTAYGAYY TNYTNTAYCA RCARTGYGAY GCNCARCCNG GNGCNWSNGG NWSNGGNGTN    960
TAYGTNMGNA TGTGGAARMG NCARCARCAR AARTGGGARM GNAARATHAT HGGNATHTTY   1020
WSNGGNCAYC ARTGGGTNGA YATGAAYGGN WSNCCNCARG AYTTYAAYGT NGCNGTNMGN   1080
ATHACNCCNY TNAARTAYGC NCARATHTGY TAYTGGATHA ARGGNAAYTA YYTNGAYTGY   1140
MGNGARGGNG AYACNGTNTT YCCNCCNGGN WSNAAY                             1176
```